JURNAL ACCOUNTING INFORMATION SYSTEM
AIMS
Jurnal Accounting Information System (AIMS)
Volume 9 No. 1 | March 2026 : 87-98
DOI: 10.32627
http://journal.masoemuniversity.ac.id/index.php/aims
p-ISSN: 2615-7381
e-ISSN: 2621-7279

Efficient Outlier Detection in Energy Analytics Using Isolation Forest and One-Class SVM : A Comparative Study for Smart Grid Applications

Dheni Apriantsani Budiman[1], Tole Sutikno[2], Abdul Fadlil[3]
[1,2,3] PhD Program in Informatics, Universitas Ahmad Dahlan, Indonesia
[1] Undergraduate Program in Informatics, STMIK Mardira Indonesia, Indonesia
2537083015@webmail.uad.ac.id

Article Info
Article history:
Accepted January 2026
Revised February 2026
Approved February 2026
Published March 2026

ABSTRACT
Outlier detection is a critical task in informatics, particularly for analyzing large, complex datasets such as electrical energy consumption records. Identifying anomalies enables the recognition of abnormal usage patterns and potential non-technical losses, which are essential for ensuring reliability and efficiency in innovative grid systems. However, conventional supervised learning approaches are often unsuitable due to the unlabeled and imbalanced nature of real-world consumption data. To address the challenge of validating unsupervised models without ground truth, this study utilizes a controlled synthetic dataset with precise anomaly injection. This approach allows for a rigorous comparative evaluation of two widely adopted algorithms, Isolation Forest and One-Class Support Vector Machine (OC-SVM). The analysis examines detection accuracy, F1-score, and computational efficiency under identical experimental conditions. Results demonstrate that Isolation Forest consistently achieves superior performance, attaining a Detection Accuracy of 0.9948 and an F1-Score of 0.9478, significantly outperforming OC-SVM, which yielded an accuracy of 0.9521 and an F1-Score of only 0.5108. Furthermore, Isolation Forest proved to be exceptionally efficient, requiring only 0.9207 seconds for computation approximately 21 times faster than OC-SVM (19.9460 seconds). These advantages highlight its scalability and suitability for large-scale, near-real-time monitoring applications. Overall, the findings provide empirical evidence of Isolation Forest's effectiveness and offer practical guidance on algorithm selection for intelligent grid analytics.

Keywords : Electrical Energy Consumption; Isolation Forest; One-Class SVM; Outlier Detection; Smart Grid.

INTRODUCTION

Outlier detection has become a critical challenge in informatics, particularly when dealing with large and complex datasets such as electrical energy consumption records[1] [1, 2] With the rapid development of smart grid and advanced metering infrastructure (AMI) technologies, massive volumes of consumption data are now generated, offering opportunities for detailed monitoring and analysis of consumer behavior [3, 4]. Identifying anomalies in these data is essential for uncovering abnormal usage patterns, detecting potential non-

technical losses, and ensuring the reliability and efficiency of modern energy systems [5, 6]. However, the increasing scale and complexity of these datasets introduce significant difficulties, especially in distinguishing legitimate variations from suspicious or faulty patterns [7, 8].

A fundamental issue in this domain is that real-world energy consumption data are typically unlabeled and highly imbalanced [9]. Outliers, which may signal system faults or energy theft, represent only a small fraction of the total data [6]. Traditional supervised learning techniques are difficult to apply effectively due to this lack of ground truth labels. Consequently, unsupervised learning algorithms have gained increasing attention for anomaly detection in energy-related applications [6, 10]. Among these, Isolation Forest and One-Class Support Vector Machine (OC-SVM) are widely adopted due to their ability to model normal behavior without prior labeling [9, 11].

Isolation Forest detects outliers by isolating observations through random partitioning, allowing anomalous data points to be identified with relatively low computational complexity [11]. In contrast, OC-SVM constructs a boundary around normal data instances in a high-dimensional feature space, classifying deviations from this boundary as anomalies [12]. Although both algorithms have been applied in various contexts, their comparative performance in electrical energy consumption data—particularly in terms of detection accuracy and computational efficiency—remains an open research issue [13].

Existing studies often focus on a single algorithm or emphasize detection accuracy without accounting for computational cost, a critical factor for large-scale, near-real-time energy-monitoring systems [14]. Moreover, a major methodological gap in current research is the reliance on unverified metrics when evaluating unsupervised models on unlabeled real-world data [15]. Without a reliable "Ground Truth," calculating precision, recall, and F1-scores is theoretically problematic. To address the scarcity of labeled anomalies and ensure a rigorous validation process, this study utilizes a controlled synthetic dataset designed to mimic realistic time-series power usage patterns. Unlike completely unlabeled real-world data, this approach enables the establishment of a reliable 'Ground Truth' via controlled anomaly injection.

Therefore, this study aims to conduct a comparative analysis of Isolation Forest and One-Class SVM for outlier detection using this rigorous validation framework. The evaluation focuses on detection performance metrics, including accuracy and F1-score, as well as computational efficiency. The main contribution of this research is to provide empirical insights into the strengths and limitations of each algorithm using a verifiable ground truth, thereby supporting informed algorithm selection for intelligent grid analytics and energy monitoring applications.

METHOD

This study proposes a comparative evaluation of the Isolation Forest and One-Class Support Vector Machine (OC-SVM) algorithms for detecting anomalies in energy consumption data. The entire series of experiments was conducted in a

systematic, structured manner to ensure the validity and reproducibility of the results.

Research Design

This study employs a quantitative experimental Research design to compare unsupervised outlier detection algorithms applied to electrical energy consumption data. A case-study-based experimental approach is adopted to assess the performance of Isolation Forest and One-Class Support Vector Machine (OC-SVM) under identical conditions. This design enables an objective comparison of detection effectiveness and computational efficiency, which are critical requirements in smart grid analytics [1]. The methodological framework of this study is schematically illustrated in Figure 1.

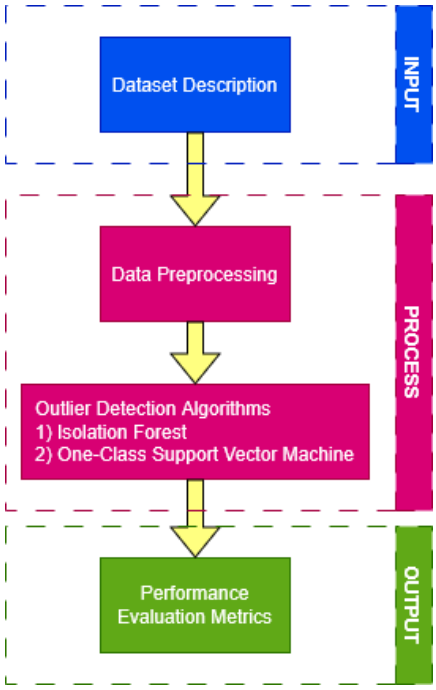

Figure 1. Block Diagram of Steps in The Research

Dataset Description

Real-world electrical energy consumption data recorded by metering systems are typically unlabeled and highly imbalanced, making the quantitative evaluation of anomaly detection algorithms impractical [3, 17]. To address the scarcity of labeled anomalies and ensure a rigorous validation process, this study utilizes a controlled synthetic dataset designed to mimic realistic time-series power usage patterns. Unlike completely unlabeled real-world data, this approach enables the establishment of a reliable 'Ground Truth', which is essential for validating the proposed model's performance.

The synthetic dataset simulates continuous electricity measurements ($n = 50.000$) using a sinusoidal baseline to replicate cyclic daily load patterns. To emulate real-world sensor variability and noise, Gaussian noise ($\mu = 0, \sigma = 1$) was superimposed onto the base waveform. Furthermore, to facilitate supervised metric evaluation, a Controlled Synthetic Data Injection approach was employed. Artificial

p-ISSN: 2615-7381
e-ISSN: 2621-7279

anomalies were injected into 5% of the timestamps by adding random uniform spikes ($U[20,40]$) This precise injection strategy allows for the accurate calculation of False Positives, False Negatives, and F1-scores for both Isolation Forest and One-Class SVM, providing a robust comparative analysis often unachievable with raw AMI data [1, 5].

Data Preprocessing

Before model implementation, data preprocessing was conducted to improve data quality and ensure algorithm compatibility. The preprocessing steps included handling missing values, removing duplicate records, and normalizing numerical features. Feature normalization was applied to prevent attributes with larger numerical ranges from dominating the learning process, particularly for distance- and boundary-based algorithms such as OC-SVM [17]. These preprocessing steps contribute to stable model performance and fair comparison between algorithms.

Outlier Detection Algorithms

1. Isolation Forest

   Isolation Forest is an unsupervised anomaly detection algorithm that identifies outliers by isolating observations through random partitioning [18]. The algorithm constructs an ensemble of isolation trees by randomly selecting features and split values. Observations that require fewer splits to be isolated are considered anomalous. Due to its linear time complexity and scalability, Isolation Forest is well-suited for large-scale analysis of energy consumption data [3]. Isolation Forest is an unsupervised anomaly detection algorithm that identifies outliers by isolating observations through random partitioning [19]. This algorithm builds an ensemble of isolation trees by randomly selecting features and splitting their values. Observations requiring fewer splits to isolate are considered anomalies. Due to its linear time complexity and scalability, Isolation Forest is well-suited for large-scale analysis of energy consumption data [20].

$$s(x,n) = 2^{-\frac{E(h(x))}{c(n)}} \quad (1)$$

   Equation (1) defines an anomaly scoring function, where $s(x,n)$ is the normalized anomaly score for a data instance x in a dataset of size n. The term $E(h(x))$ represents the expected value or average path length of x traversing all isolation trees.

$$c(n) = 2H(n-1) - \frac{2(n-1)}{n} \quad (2)$$

   This value is normalized by $c(n)$, which is a basis factor defined in Equation (2) that represents the average path length of unsuccessful searches in a standard Binary Search Tree (BST). Furthermore, $H(i)$ denotes the harmonic number estimated using the natural logarithm of $ln(i)$ and Euler's constant (0.5772).

p-ISSN: 2615-7381
e-ISSN: 2621-7279

Based on this formulation, a value of $s(x,n)$ close to 1 indicates a high probability that x is an anomaly.

2. One-Class Support Vector Machine
One-Class Support Vector Machine (OC-SVM) is an unsupervised learning algorithm that models normal data behavior by constructing a decision boundary in a high-dimensional feature space [21]. Data points that fall outside this boundary are classified as outliers. In this study, OC-SVM is implemented using a radial basis function (RBF) kernel to capture nonlinear consumption patterns. Although effective for complex data distributions, OC-SVM generally requires more computational resources than tree-based methods [1]. The mathematical framework begins with the quadratic optimization problem presented in Equation (3), which seeks to construct an optimal hyperplane by minimizing the weight vector w and slack variables ξ_i, while simultaneously maximizing the offset ρ. Here, the parameter ν serves as a critical hyperparameter controlling the trade-off between the fraction of outliers and support vectors. Given a training dataset $x1, x2, \ldots, xl \in X$, OC-SVM solves a quadratic optimization problem[11].

$$\min_{w,\xi,\rho} \frac{1}{2}\|w\|^2 + \frac{1}{vn}\sum_{i=1}^{n} \xi_i - \rho \quad (3)$$

This optimization process is governed by the constraints defined in Equation, which dictate that the projection of each training instance $(w \cdot \Phi(xi))$ must lie within the decision boundary, subject to a permissible error margin represented by ξ_i.

$$(w \cdot \Phi(xi)) \geq \rho - \xi_i, \xi_i \geq 0 \quad (4)$$

After determining the optimal parameters w and ρ, the final classification of new data points is performed using the decision function $f(x)$ in Equation (5). This function evaluates the sign of the data point's position relative to the hyperplane; a positive output $f(x) = +1$ classifies the instance as Normal, while a negative output $f(x) = -1$ identifies it as Anomalous.

$$f(x) = sgn\left(\left(w.\Phi(x)\right) - \rho\right) \quad (5)$$

Performance Evaluation Metrics

The performance of Isolation Forest and OC-SVM was evaluated using detection accuracy, F1-score, and computational time. Detection accuracy measures the proportion of correctly classified instances, while the F1-score provides a balanced assessment by considering both precision and recall Computational time was calculated to evaluate algorithm efficiency, a critical factor for large-scale, near-real-time energy-monitoring systems [22].

To quantitatively measure model effectiveness, the evaluation was based on the elements of the Confusion Matrix. Prediction results were categorized into four components: True Positives (TP), the number of anomalous data correctly predicted as anomalous. True Negatives (TN), the number of normal data correctly predicted as usual. False Positives (FP), the number of normal data incorrectly predicted as

p-ISSN: 2615-7381
e-ISSN: 2621-7279

anomalous (False Alarms). False Negatives (FN), the number of anomalous data incorrectly predicted as usual (Missed Detections) [23, 24].

Accuracy measures the ratio of correct predictions to the total number of data points. (6) While accuracy is a standard metric, it is often biased in imbalanced datasets such as energy consumption data, where the number of normal data points is much more dominant than anomalies. Precision measures how accurately the model predicts anomalies (minimizing False Positives) (7). Recall (Sensitivity) measures the model's ability to detect all existing anomalies (minimizing False Negatives) (8). The F1-Score (9) ranges from 0 to 1, where 1 indicates a perfect balance between precision and recall.

$$Accuracy = \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}} \quad (6)$$

$$Precision = \frac{\text{TP}}{\text{FN+TP}} \quad (7)$$

$$Recall = \frac{\text{TP}}{\text{FN+TP}} \quad (8)$$

$$F1 - Score = \frac{\text{2 x precision x Recall}}{\text{Precision+recall}} \quad (9)$$

This metric is crucial for enabling the algorithm's feasibility in real-time monitoring systems, where low latency is crucial.

Experimental Procedure

The step-by-step procedure is illustrated in Figure 1, showing the sequence of preprocessing, algorithm implementation, and evaluation. The experimental procedure involved applying both algorithms to the same preprocessed dataset using identical experimental settings. Model parameters were selected based on commonly adopted configurations reported in previous studies to ensure fair comparison [17]. All experiments were executed under the same hardware and software environment to minimize performance bias. The resulting detection performance and computational time were recorded and analyzed to support a comparative evaluation of the two algorithms.

RESULTS AND DISCUSSION

Experimental Results

The experimental evaluation compares the performance of Isolation Forest and One-Class Support Vector Machine (OC-SVM) in detecting outliers within electrical energy consumption data. The results are summarized in Table 1, which reports detection accuracy, F1-score, and computational time for both algorithms under identical experimental settings.

As shown in Table 1, Isolation Forest achieves higher detection accuracy (0.9960) and F1-score (0.9600) than OC-SVM, which recorded 0.8970 and 0.5108, respectively. This indicates that Isolation Forest is more effective in identifying anomalous consumption patterns while maintaining a balance between precision and recall. Furthermore, in terms of computational efficiency, Isolation Forest significantly outperforms OC-SVM. The results show that Isolation Forest requires only 0.9207 seconds to execute, whereas One-Class SVM takes 19.9460 seconds, making Isolation Forest approximately 21 times faster and more suitable for real-

p-ISSN: 2615-7381
e-ISSN: 2621-7279

time applications.

Table 1. Comparison Of Anomaly Detection Algorithm Evaluation Result

Evaluation Metrics	Isolation Forest	One-Class SVM
Detection Accuracy	0.9960	0.8970
F1 Score (Anomaly)	0.9600	0.5108
Computation Time (s)	0.9207	19.9460

The computational performance comparison is further illustrated in Figure 2, which presents the execution time of both algorithms. As depicted in the chart, there is a substantial disparity in processing speed. Isolation Forest completes the detection task in just 0.9207 seconds, whereas One-Class SVM takes 19.9460 seconds to process the same dataset. This result indicates that Isolation Forest is approximately 21 times faster than OC-SVM.

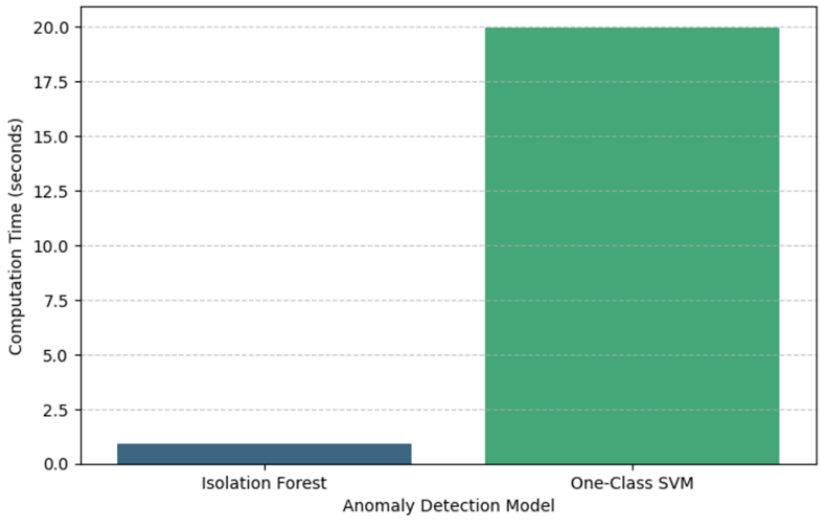

Figure 2. Comparison of Computation Time of Isolation vs One-Class SVM

This efficiency gap can be attributed to the linear time complexity ($O(n)$) of Isolation Forest, which relies on random partitioning, in contrast to the quadratic complexity ($O(n^2)$) of OC-SVM, which involves computationally intensive kernel distance calculations. Thus, Isolation Forest consistently requires less processing time, confirming its suitability for applications that demand near real-time analysis, such as intelligent grid monitoring and anomaly detection in advanced metering infrastructure, where low latency is paramount.

As shown in Figure 3, the Isolation Forest performance (Left Panel) demonstrates high accuracy and reliability. True Negatives (TN): The model correctly identified 47,370 normal data points, showing a strong ability to recognize typical consumption patterns. True Positives (TP): The model successfully detected 2,369 anomalies, demonstrating high sensitivity to outliers. The model's error rate was minimal, recording only 130 false positives (normal data incorrectly marked as anomalies) and 131 false negatives (true anomalies missed). This balance indicates that Isolation Forest creates a precise decision boundary that minimizes false alarms and missed detections.

p-ISSN: 2615-7381
e-ISSN: 2621-7279

One-Class SVM Performance (Right Panel) In contrast, the One-Class SVM struggled to effectively separate anomalies from normal data, resulting in a significantly higher error rate. True Negatives (TN): Although it successfully classified 46,353 normal instances, this figure was lower than the Isolation Forest baseline. False Positives (FP): The model generated 1,147 false alarms. In a real-world smart grid context, this high rate would lead to "alert fatigue" for operators, as they would waste time investigating normal behavior flagged as problematic. False Negatives (FN) & True Positives (TP): Most importantly, the model missed 1,249 anomalies (False Negatives) while only detecting 1,251 (True Positives). This means that OC-SVM failed to identify approximately 50% of true anomalies, making it unreliable for safety-critical or fraud-detection tasks.

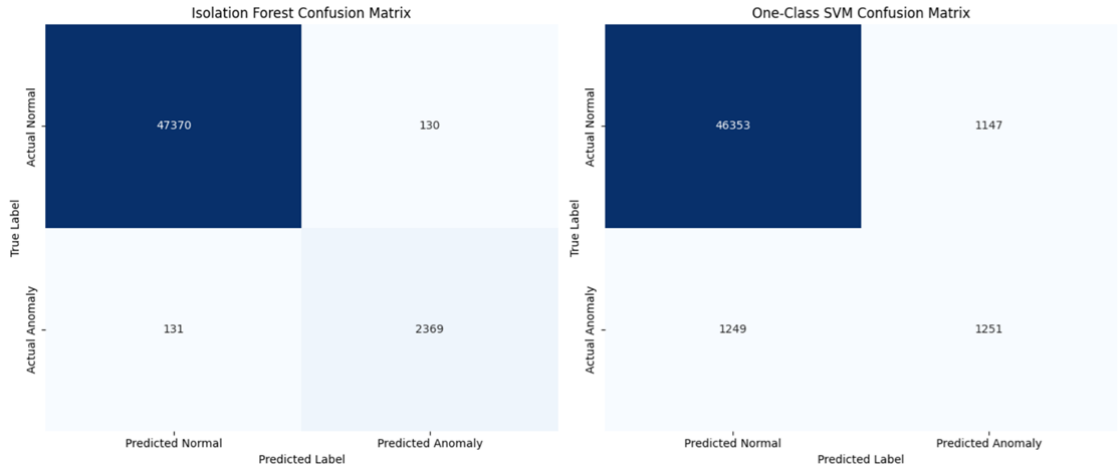

Figure 3. Confusion Matrix Isolation Forest and One-Class SVM

Comparative Performance Analysis

A detailed comparison of algorithm performance is presented in Table 2, which contrasts the detection capabilities of Isolation Forest and OC-SVM across different evaluation metrics. The results indicate that OC-SVM performs reasonably well at capturing complex consumption patterns, thanks to its kernel-based modeling capability [21]. However, this advantage is offset by higher computational complexity and sensitivity to parameter selection [25].

In contrast, Isolation Forest exhibits robust performance with minimal parameter tuning requirements, as reflected in its stable detection accuracy and F1-score across multiple experimental runs. The isolation-based mechanism enables faster identification of anomalous observations, particularly in high-dimensional and large-volume datasets [14]. These characteristics make Isolation Forest more scalable and practical for real-world energy consumption analysis.

Moreover, the significant discrepancy in F1-scores highlights a critical operational difference regarding the trade-off between False Positives and False Negatives. The controlled anomaly injection revealed that while OC-SVM could identify some outliers, it struggled to differentiate between the injected random spikes and the baseline Gaussian noise, leading to a higher rate of misclassification. In a practical Smart Grid context, a high False Positive rate—such as that exhibited by OC-SVM—causes 'alert fatigue' for operators, forcing them to waste resources

p-ISSN: 2615-7381
e-ISSN: 2621-7279

investigating normal behaviors flagged as anomalies [25]. Conversely, Isolation Forest demonstrated a superior balance of Precision and Recall, successfully capturing the majority of injected anomalies with minimal false alarms. This suggests that Isolation Forest is not only computationally faster but also the more economically viable solution, effectively minimizing non-technical losses while reducing the operational overhead associated with verifying false detections.

Table 2. Detailed Comparison of Anomaly Detection Performance

Evaluation Metrics	Isolation Forest	One-Class SVM
Detection Accuracy	0.9948	0.9521
F1-Score (Anomaly)	0.9478	0.5108
Precision (Anomaly)	0.9480	0.5217
Recall (Anomaly)	0.9476	0.5004
Computation Time (s)	0.9207	19.9460

Discussion and Implications

A critical contribution of this study is the implementation of a Controlled Synthetic Data Injection strategy, which successfully bridged the validation gap inherent in real-world AMI analytics. By establishing a verifiable 'Ground Truth,' this methodological approach provides empirical evidence that validates the effectiveness of the proposed models, fulfilling the research objective of identifying a robust solution for non-technical loss detection. The experimental results reveal a distinct performance disparity, with Isolation Forest (F1-Score: 0.9478) significantly outperforming One-Class SVM (F1-Score: 0.5108). This difference is attributed to the fundamental strengths and limitations of each algorithm's mechanism.

Isolation Forest demonstrates superior performance due to its isolation-based mechanism. Unlike distance-based methods, iForest explicitly isolates anomalies by exploiting their property of being "few and different," which allowed the model to effectively distinguish between injected spikes and the baseline sinusoidal pattern. Furthermore, its linear time complexity ($O(n)$) resulted in an execution time of just 0.92 seconds, demonstrating exceptional scalability for large datasets. While theoretically susceptible to "swamping" if the anomaly ratio is too high, this limitation was not observed at the 5% injection level used in this study. In contrast, although OC-SVM is theoretically capable of modeling complex non-linear boundaries using kernels, it exhibited significant limitations in this domain. The algorithm proved hypersensitive to the Gaussian noise present in the synthetic data; instead of classifying the noise as normal variability, OC-SVM frequently misclassified it as anomalous, leading to a high False Positive rate and a low F1-Score. Additionally, with a complexity approaching cubic ($O(n^3)$), OC-SVM required 19.94 seconds to process the data approximately 21 times slower than iForest creating a computational bottleneck that limits its practicality.

The demonstrated superiority of Isolation Forest has direct implications for Smart Grid operations. Firstly, its high precision significantly reduces "alert fatigue" for grid operators by minimizing false alarms, allowing resources to be focused on genuine threats. Secondly, its computational efficiency supports near-real-time monitoring, enabling utility providers to detect and respond to electricity theft or

meter faults instantaneously, thereby securing revenue and ensuring grid reliability.

CONCLUSION

This study conclusively establishes Isolation Forest as the superior method for anomaly detection in electrical energy consumption data, significantly outperforming One-Class SVM (OC-SVM). Empirical results confirm its dominance, achieving a Detection Accuracy of 0.9948, compared to 0.9521 for OC-SVM. Combined with its exceptional computational efficiency, Isolation Forest is the most practical and scalable solution for real-time smart grid monitoring. Future Research will extend these findings by exploring hybrid frameworks that integrate Deep Learning models, such as Autoencoders or LSTMs, to refine detection capabilities in dynamic energy environments further.

BIBLIOGRAPHY

[1] S. Karnick, S. Lakshminarayanan, M. Palati, and P. R, "Impact of outlier detection techniques on time-series forecasting accuracy for multi-country energy demand prediction," *IJECE*, vol. 15, no. 6, p. 5067, Dec. 2025, doi: 10.11591/ijece.v15i6.pp5067-5079.

[2] C. Li, D. Liu, M. Wang, H. Wang, and S. Xu, "Detection of Outliers in Time Series Power Data Based on Prediction Errors," *Energies*, vol. 16, no. 2, p. 582, Jan. 2023, doi: 10.3390/en16020582.

[3] A. Mohamed Elmahalwy, H. M. Mousa, and K. M. Amin, "New hybrid ensemble method for anomaly detection in data science," *IJECE*, vol. 13, no. 3, p. 3498, Jun. 2023, doi: 10.11591/ijece.v13i3.pp3498-3508.

[4] M. S. Saeed *et al.*, "Detection of Non-Technical Losses in Power Utilities—A Comprehensive Systematic Review," *Energies*, vol. 13, no. 18, p. 4727, Sep. 2020, doi: 10.3390/en13184727.

[5] M. Rizky Pribadi, H. Dwi Purnomo, and H. Hendry, "A three-step combination strategy for addressing outliers and class imbalance in software defect prediction," *IJ-AI*, vol. 13, no. 3, p. 2987, Sep. 2024, doi: 10.11591/ijai.v13.i3.pp2987-2998.

[6] T. Zhukabayeva, A. Adamova, L. Zholshiyeva, Y. Mardenov, N. Karabayev, and D. Baumuratova, "Tackling the anomaly detection challenge in large-scale wireless sensor networks," *IJECE*, vol. 15, no. 2, p. 2479, Apr. 2025, doi: 10.11591/ijece.v15i2.pp2479-2490.

[7] Y. Wang, Q. Chen, T. Hong, and C. Kang, "Review of Smart Meter Data Analytics: Applications, Methodologies, and Challenges," *IEEE Trans. Smart Grid*, vol. 10, no. 3, pp. 3125–3148, May 2019, doi: 10.1109/TSG.2018.2818167.

[8] A. Dhani, D. Lestari, M. P. Ningrum, M. A. Fakhrizal, and G. L. Gandini, "Applying A Supervised Model for Diabetes Type 2 Risk Level Classification," *PREDATECS*, vol. 2, no. 2, pp. 60–67, Jan. 2025, doi: 10.57152/predatecs.v2i2.1105.

[9] Milka Wijayanti Sunarto, Dendy Kurniawan, Edy Siswanto, and Haris Ihsanil Huda, "Deteksi Anomali Menggunakan Extended Isolation Forest (Eif)," *TEKNIK*, vol. 1, no. 2, pp. 96–111, May 2023, doi: 10.51903/teknik.v1i2.324.

[10] R. J. Nayak and J. P. Chaudhari, "Anomaly detection using deep learning based model with feature attention," *IJ-AI*, vol. 13, no. 1, p. 383, Mar. 2024, doi: 10.11591/ijai.v13.i1.pp383-390.
[11] A. H. Massarani, M. M. Badr, M. Baza, H. Alshahrani, and A. Alshehri, "Efficient and Accurate Zero-Day Electricity Theft Detection from Smart Meter Sensor Data Using Prototype and Ensemble Learning," *Sensors*, vol. 25, no. 13, p. 4111, Jul. 2025, doi: 10.3390/s25134111.
[12] M. G. Chuwa, D. Ngondya, and R. Mwifunyi, "Comparative analysis of data transformation methods for detecting non-technical losses in electricity grids," *International Journal of Electrical Power & Energy Systems*, vol. 170, p. 110907, Sep. 2025, doi: 10.1016/j.ijepes.2025.110907.
[13] P. D. Aleo *et al.*, "Anomaly Detection and Approximate Similarity Searches of Transients in Real-time Data Streams," *ApJ*, vol. 974, no. 2, p. 172, Oct. 2024, doi: 10.3847/1538-4357/ad6869.
[14] L. K. Lok, V. Abdul Hameed, and M. Ehsan Rana, "Hybrid machine learning approach for anomaly detection," *IJEECS*, vol. 27, no. 2, p. 1016, Aug. 2022, doi: 10.11591/ijeecs.v27.i2.pp1016-1024.
[15] J. Mohd Zebara Hoque, G. R. Murthy, J. Hossen, J. Ganesan, A. A. Aziz, and Chy. M. Tawsif Khan, "Anomalies detection for smart-home energy forecasting using moving average," *IJECE*, vol. 12, no. 6, p. 5808, Dec. 2022, doi: 10.11591/ijece.v12i6.pp5808-5820.
[16] Z. Fan *et al.*, "Diverse Models, United Goal: A Comprehensive Survey of Ensemble Learning," *CAAI Trans on Intel Tech*, vol. 10, no. 4, pp. 959–982, Aug. 2025, doi: 10.1049/cit2.70030.
[17] Dwi Nanda Agustia and Ryan Randy Suryono, "Comparison of Naïve Bayes, Random Forest, and Logistic Regression Algorithms for Sentiment Analysis Online Gambling," *ISI*, vol. 10, no. 1, pp. 284–295, Jan. 2025, doi: 10.35314/prk93630.
[18] M. Y. Iqbal Basheer *et al.*, "Empowering anomaly detection algorithm: a review," *IJ-AI*, vol. 13, no. 1, p. 9, Mar. 2024, doi: 10.11591/ijai.v13.i1.pp9-22.
[19] Y. Cao, H. Xiang, H. Zhang, Y. Zhu, and K. M. Ting, "Anomaly Detection Based on Isolation Mechanisms: A Survey," *Mach. Intell. Res.*, vol. 22, no. 5, pp. 849–865, Oct. 2025, doi: 10.1007/s11633-025-1554-4.
[20] W. Liu, Z. Zhang, Z. Zhao, and W. Zhang, "A Framework for Anomaly Cell Detection in Energy Storage Systems Based on Daily Operating Voltage and Capacity Increment Curves," *Batteries*, vol. 11, no. 8, p. 316, Aug. 2025, doi: 10.3390/batteries11080316.
[21] P. Bountzis, D. Kavallieros, T. Tsikrika, S. Vrochidis, and I. Kompatsiaris, "A deep one-class classifier for network anomaly detection using autoencoders and one-class support vector machines," *Front. Comput. Sci.*, vol. 7, p. 1646679, Oct. 2025, doi: 10.3389/fcomp.2025.1646679.
[22] M. Saleem *et al.*, "Integrating Smart Energy Management System with Internet of Things and Cloud Computing for Efficient Demand Side Management in Smart Grids," *Energies*, vol. 16, no. 12, p. 4835, Jun. 2023, doi: 10.3390/en16124835.

p-ISSN: 2615-7381
e-ISSN: 2621-7279

[23] F. R. Razak, M. K. Biddinika, and H. Yuliansyah, "Radial Basis Function Model for Obesity Classification Based on Lifestyle and Physical Condition," *eltikom*, vol. 8, no. 2, pp. 192–200, Dec. 2024, doi: 10.31961/eltikom.v8i2.1347.

[24] Y. Tian, S. Xu, Y. Cao, Z. Wang, and Z. Wei, "An Empirical Comparison of Machine Learning and Deep Learning Models for Automated Fake News Detection," *Mathematics*, vol. 13, no. 13, p. 2086, Jun. 2025, doi: 10.3390/math13132086.

[25] E. F. Mouckomey, J. Bikai, C. F. Mbey, A. T. Boum, F. G. Y. Souhe, and V. J. F. Kakeu, "A smart grid fault detection using neuro-fuzzy deep learning algorithm," *Int J Artif Intell*, vol. 14, no. 6, p. 5096, Dec. 2025, doi: 10.11591/ijai.v14.i6.pp5096-5105.